Kinetic hard-modelling and spectral validation of rank-deficient spectroscopic data

11th Scandinavian Symposium on Chemometrics (SSC11)
Loen/Stryn – Norway, June 9th 2009

Julien Billeter

julien.billeter@chem.ethz.ch

ETH Zürich, Institute for Chemical and Bioengineering,
Safety and Environmental Technology Group, Switzerland

ETH
Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

Safety and Environmental Technology Group

Direct fitting by modelling concentrations

$$\boxed{\mathbf{Y} = \mathbf{C}\,\mathbf{A} + \mathbf{R}}$$

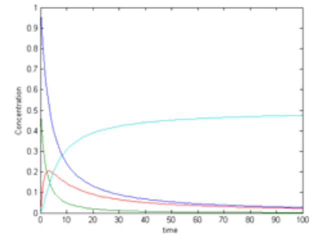
Beer's law

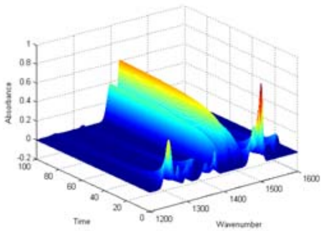
Direct fitting by modelling concentrations

Modelled concentrations

$\mathbf{C}_{\text{modelled}} = \text{function}\ (\mathbf{k})$

(*nt* x *ns*)





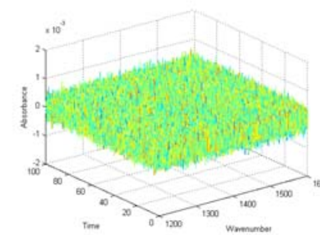
(*nt* x *nw*)

Y: Measured absorbance

$$\mathbf{Y} = \mathbf{C}\,\mathbf{A} + \mathbf{R}$$

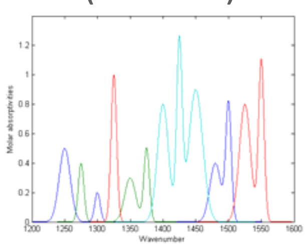
Beer's law

R: Residuals



(*nt* x *nw*)

(*ns* x *nw*)

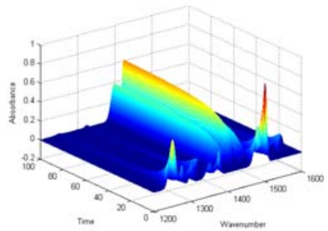


A: Pure component spectra

Implicit calibration $\mathbf{A} = \mathbf{C}^{+}_{\text{modelled}}\mathbf{Y}$

Explicit calibration $\mathbf{A} = (\mathbf{C}^{+}\mathbf{Y})_{\text{calibration}}$

Direct fitting by modelling concentrations

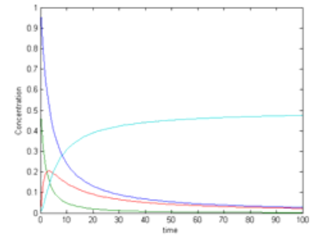


(nt x nw)

Modelled concentrations

$\mathbf{C}_{\text{modelled}} = \text{function}\ (\mathbf{k})$

(nt x ns)

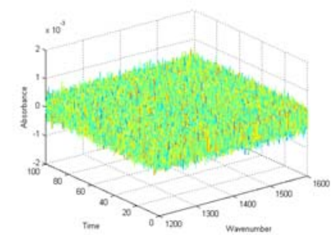


Y: Measured absorbance

$$\mathbf{Y} = \mathbf{C}\,\mathbf{A} + \mathbf{R}$$

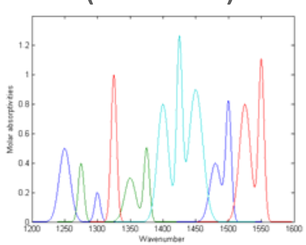
Beer's law

R: Residuals



(nt x nw)

(ns x nw)



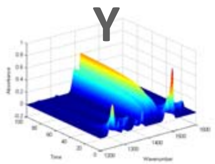
A: Pure component spectra

Implicit calibration $\mathbf{A} = \mathbf{C}^{+}_{\text{modelled}}\mathbf{Y}$

Explicit calibration $\mathbf{A} = (\mathbf{C}^{+}\mathbf{Y})_{\text{calibration}}$

Least squares fitting $\min\limits_{\mathbf{k}} \left\| \mathbf{Y} - \mathbf{C}_{\text{modelled}}\mathbf{A} \right\|^2$

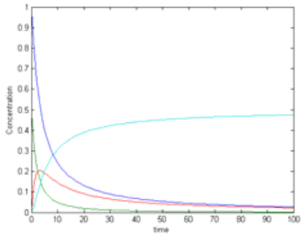
Rank deficiency in spectroscopy



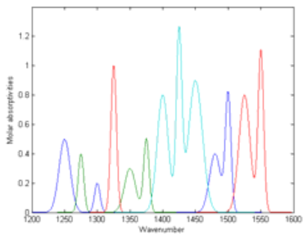
$$\begin{cases} A + B \xrightarrow{k_1} C \\ A + C \xrightarrow{k_2} D \end{cases}$$

Rank deficiency in spectroscopy

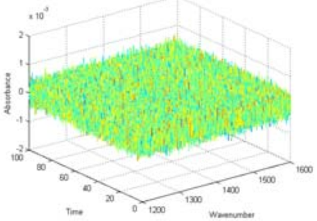
C



A



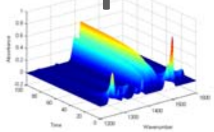
R



Beer's law $\left(\mathbf{Y} = \mathbf{C}\ \mathbf{A}\right)$

ns = 4 species

Y



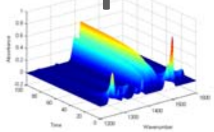
$$\begin{cases} A + B \xrightarrow{k_1} C \\ A + C \xrightarrow{k_2} D \end{cases}$$

Rank deficiency in spectroscopy

Beer's law $\left(\mathbf{Y} = \mathbf{C}\ \mathbf{A}\right)$

ns = 4 species

Y

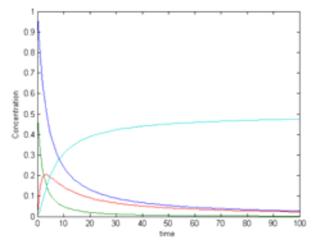


$$\begin{cases} A + B \xrightarrow{k_1} C \\ A + C \xrightarrow{k_2} D \end{cases}$$

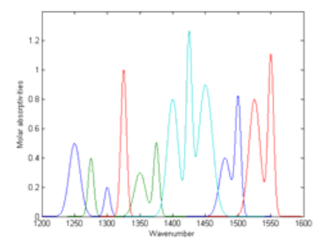
PCA / TFA $\left(\overline{\mathbf{Y}} = \overline{\mathbf{U}}\ \overline{\mathbf{S}}\ \overline{\mathbf{V}}\right)$

nc = 3 factors

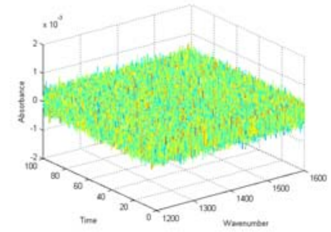
C



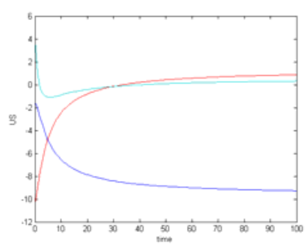
A



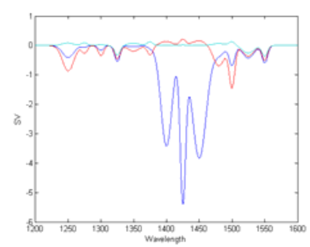
R



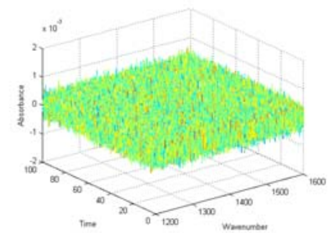
$\overline{\mathbf{C}} = \overline{\mathbf{U}}\ \mathbf{T}$



$\overline{\mathbf{A}} = \mathbf{T}^{-1}\overline{\mathbf{S}}\ \overline{\mathbf{V}}$



R

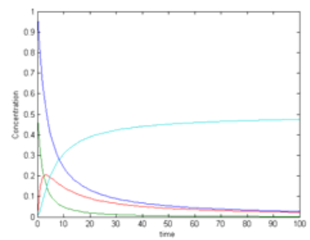


Where **T** is a transformation matrix of dimensions (*nc* x *nc*)

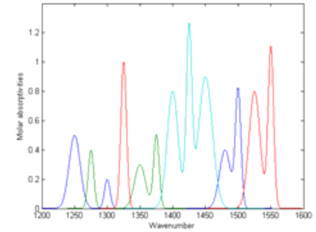
Rank deficiency in spectroscopy

Beer's law $\left(\mathbf{Y} = \mathbf{C}\ \mathbf{A}\right)$
ns = 4 species

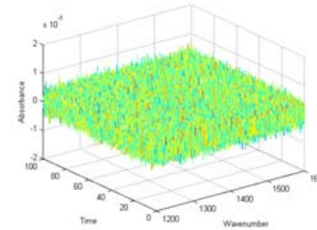
C



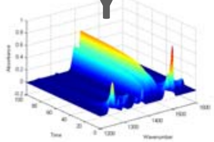
A



R



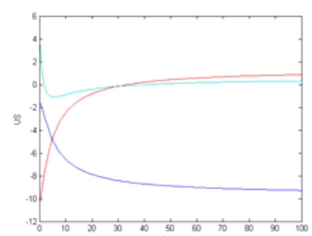
Y



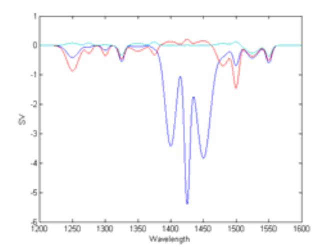
$$\begin{cases} A + B \xrightarrow{k_1} C \\ A + C \xrightarrow{k_2} D \end{cases}$$

PCA / TFA $\left(\bar{\mathbf{Y}} = \bar{\mathbf{U}}\ \bar{\mathbf{S}}\ \bar{\mathbf{V}}\right)$
nc = 3 factors

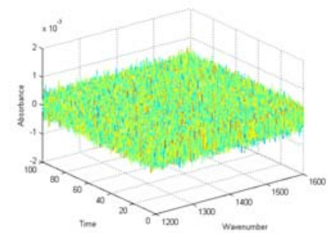
$\bar{\mathbf{C}} = \bar{\mathbf{U}}\ \mathbf{T}$



$\bar{\mathbf{A}} = \mathbf{T}^{-1}\bar{\mathbf{S}}\ \bar{\mathbf{V}}$



R



Where **T** is a transformation matrix of dimensions (*nc* x *nc*)

Spectroscopic data matrix **Y** is rank deficient when:

Significant factors (*nc*) in PCA < number of reactive species (*ns*) $\Leftrightarrow$ rank(**Y**) < *ns*

Sources and problems of rank deficiency

$$\mathbf{Y} = \mathbf{C}\,\mathbf{A}$$

Sources and problems of rank deficiency

$$\mathbf{Y} = \mathbf{C}\,\mathbf{A}$$

Rank deficiency in **Y** is due to

Linear dependencies in **C** and/or Linear dependencies in **A**

Mathematical ambiguity in case of <u>implicit calibration</u>

A cannot be computed by $\mathbf{C}^{+}\mathbf{Y}$ as **A** is not unique

Not discussed here

All spectra in **A** are assumed to be linearly independent

<u>Example</u>: two species that are consumed or generated at the same rate

$$\mathrm{rank}(\mathbf{Y}) = \min\left[\mathrm{rank}(\mathbf{C}),\ \mathrm{rank}(\mathbf{A})\right] = \mathrm{rank}(\mathbf{C})$$

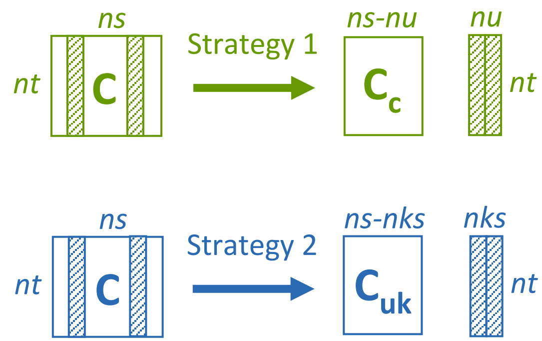
Strategies to treat rank deficiency in concentrations

Strategies to treat rank deficiency in concentrations

Model reduction

Strategy 1
define *nu* uncoloured species

Strategy 2
include *nks* known spectra in the analysis (explicit calibration)



Partial spectral resolution

Full spectral resolution

Strategies to treat rank deficiency in concentrations

Model reduction

Strategy 1

define *nu* uncoloured species

Strategy 2

include *nks* known
spectra in the analysis
(explicit calibration)

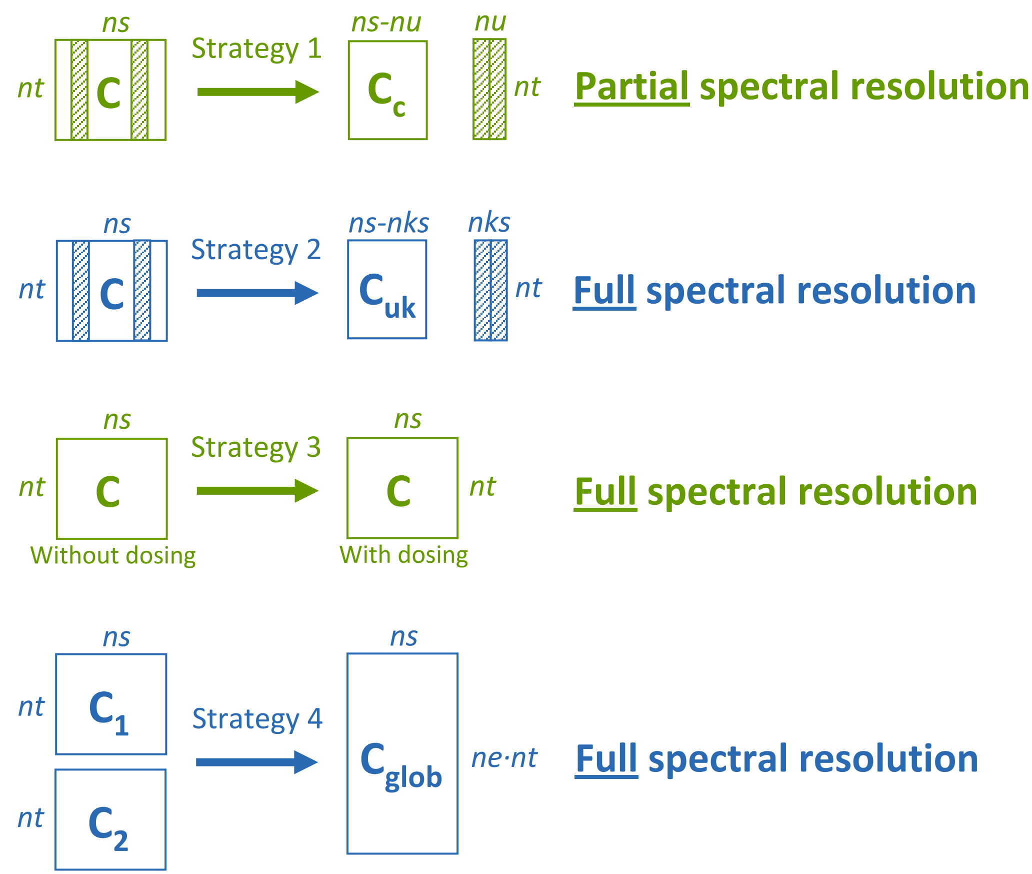
Rank augmentation

Strategy 3

dose one or more species
in *nf* dosing steps

Strategy 4:

perform *ne* additional experiments
by varying the initial concentrations
(second order global analysis)



Strategies to treat rank deficiency in concentrations

Model reduction

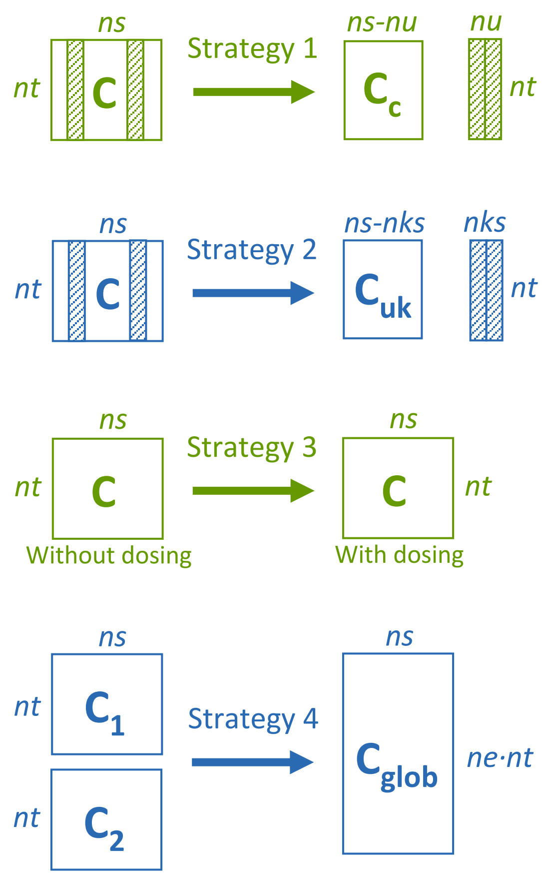
Strategy 1
define *nu* uncoloured species

Strategy 2
include *nks* known
spectra in the analysis
(explicit calibration)

Rank augmentation

Strategy 3
dose one or more species
in *nf* dosing steps

Strategy 4:
perform *ne* additional experiments
by varying the initial concentrations
(second order global analysis)



<u>Partial</u> spectral resolution

<u>Full</u> spectral resolution

<u>Full</u> spectral resolution

<u>Full</u> spectral resolution

How to identify the species to include in these four Strategies ?

Rank and kernel of the concentration matrix

Rank of **C**

Number of linearly independent columns or rows in **C**

>> Defines the maximum number of columns (species) to keep in Strategy 1 (*ns* – *nu*) and Strategy 2 (*ns* – *nks*)

kernel of **C**

Vector space spanned by the vectors forming the null space **0** when multiplied by **C**

$$\text{e.g. } \ker \mathbf{C} = \begin{array}{cc} \begin{bmatrix} 0.8 & 0.1 \\ -0.3 & -0.7 \\ 0 & 0 \\ 0.5 & -0.6 \end{bmatrix} & \begin{array}{c} A \\ B \\ C \\ D \end{array} \end{array}$$

>> ker **C** defines a mass balance equation: **C** (ker **C**) = **0**

>> ker **C** defines which columns of **C** are linearly dependent or independent

Interpretation of the kernel of C

$$\text{e.g. ker } \mathbf{C} = \begin{array}{c} \left[\begin{array}{cc} \boxed{0.8 \quad 0.1} \\ -0.3 & -0.7 \\ \boxed{0 \quad 0} \\ 0.5 & -0.6 \end{array}\right] \end{array} \begin{array}{c} \boxed{A} \\ B \\ \boxed{C} \\ D \end{array}$$

	Species without zero rows in ker C	Species with zero rows in ker C
kernel of C	Linearly dependent species	Linearly independent species

Consequences for Strategies 1 – 4 ?

Interpretation of the kernel of C

e.g. $\ker \mathbf{C} = \begin{array}{c} \begin{bmatrix} 0.8 & 0.1 \\ -0.3 & -0.7 \\ 0 & 0 \\ 0.5 & -0.6 \end{bmatrix} \end{array} \begin{array}{c} \boxed{A} \\ B \\ \boxed{C} \\ D \end{array}$

	Species without zero rows in ker C	Species with zero rows in ker C
kernel of C	**Linearly dependent species**	**Linearly independent species**

Consequences for Strategies 1 – 4

Strategy 1	Define this species as uncoloured…	Do not define this species as uncoloured…
Strategy 2	Include its pure spectrum…	Do not provide its pure spectrum…
Strategy 3	Dose this species…	Do not dose this species…
Strategy 4	Vary its initial concentration…	Do not vary its initial concentration…

… to break/avoid rank deficiency in **C**

Interpretation of the kernel of C

e.g. $\ker \mathbf{C} = \begin{array}{c} \left[\begin{array}{cc} 0.8 & 0.1 \\ -0.3 & -0.7 \\ 0 & 0 \\ 0.5 & -0.6 \end{array}\right] \end{array} \begin{array}{c} A \\ B \\ C \\ D \end{array}$

	Species without zero rows in ker C	Species with zero rows in ker C
kernel of C	**Linearly dependent species**	**Linearly independent species**
	Consequences for Strategies 1 – 4	
Strategy 1	Define this species as uncoloured…	Do not define this species as uncoloured…
Strategy 2	Include its pure spectrum…	Do not provide its pure spectrum…
Strategy 3	Dose this species…	Do not dose this species…
Strategy 4	Vary its initial concentration…	Do not vary its initial concentration…

… to break/avoid rank deficiency in **C**

Can we obtain this information without numerical integration of the rate laws, i.e. using a <u>time invariant</u> approach?

Modelling kinetic concentration profiles

Kinetic rate law

$$\dot{\mathbf{x}}_t = \prod_{\substack{\text{along}\\ \text{columns}}} \left(\mathbf{c}_t^T \mathbf{1}\right)^{\cdot \mathbf{E}^T} \mathrm{DIAG}\left(\mathbf{k}\right)$$

Reactant coefficients
$\mathbf{E}$ $(nr \text{ x } ns)$

Rate constants
$\mathbf{k}$ $(1 \text{ x } nr)$

Stoichiometric coefficients
$\mathbf{N}$ $(nr \text{ x } ns)$

Concentration profiles

$$\dot{\mathbf{c}}_t = \dot{\mathbf{x}}_t \mathbf{N} + f_{\text{in},t} V_t^{-1}\left(\mathbf{c}_{\text{in},t} - \mathbf{c}_t\right) - f_{\text{out},t} V_t^{-1} \mathbf{c}_t$$

This system of ODE is numerically integrated with initial concentrations $\mathbf{c}_0$ $(1 \text{ x } ns)$ and results in the concentration matrix $\mathbf{C}$ $(nt \text{ x } ns)$

Modelling kinetic concentration profiles

$$\begin{cases} A + B \xrightarrow{k_1} C \\ A + C \xrightarrow{k_2} D \end{cases}$$

Kinetic rate law

$$\boxed{\dot{\mathbf{x}}_t = \prod_{\substack{\text{along}\\ \text{columns}}} \left(\mathbf{c}_t^T \mathbf{1}\right)^{\bullet \mathbf{E}^T} \mathrm{DIAG}(\mathbf{k})}$$

$$\begin{bmatrix} k_1 c_{t,A} c_{t,B} & k_2 c_{t,A} c_{t,C} \end{bmatrix}$$

$$\begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \end{bmatrix}$$

Reactant coefficients
$\mathbf{E}\ (nr \text{ x } ns)$

Rate constants
$\mathbf{k}\ (1 \text{ x } nr)$ $\begin{bmatrix} k_1 & k_2 \end{bmatrix}$

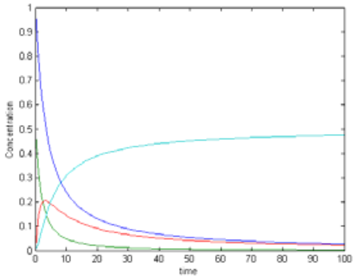
$$\begin{bmatrix} -1 & -1 & 1 & 0 \\ -1 & 0 & -1 & 0 \end{bmatrix}$$

Stoichiometric coefficients
$\mathbf{N}\ (nr \text{ x } ns)$

Concentration profiles

$$\boxed{\dot{\mathbf{c}}_t = \dot{\mathbf{x}}_t \mathbf{N} + f_{\text{in},t} v_t^{-1}\left(\mathbf{c}_{\text{in},t} - \mathbf{c}_t\right) - f_{\text{out},t} v_t^{-1} \mathbf{c}_t}$$

This system of ODE is numerically integrated with initial concentrations $\mathbf{c_0}$ $(1 \text{ x } ns)$ and results in the concentration matrix $\mathbf{C}$ $(nt \text{ x } ns)$



A time invariant matrix Ω equivalent to C

$$\boldsymbol{\Omega} = \left[\begin{array}{c} (\mu \mathbf{1})^{\bullet \mathbf{E}^T} \mathrm{DIAG}(\mathbf{k})\mathbf{N} \\ \hline \mathbf{c_0} \\ \hline \mathbf{C_{in}} \\ \hline \mathbf{C_0^{ne}} \end{array} \right] \quad (ns + 1 + nf + ne \ \times \ ns - nu - nks) \qquad \ker \boldsymbol{\Omega} = (\ker \mathbf{C})\mathbf{T_{lin}}$$

μ	an arbitrary positive scalar different from 0 and 1
1 (*ns* x *nr*)	matrix comprised of ones
E (*nr* x *ns*)	matrix of reactant coefficients
$\bullet \mathbf{E}^T$	element-wise raise to the power of $\mathbf{E}^T$
DIAG	operator generating a diagonal matrix from a vector argument
k (1 x *nr*)	vector of rate constants
N (*nr* x *ns*)	matrix of stoichiometric coefficients
$\mathbf{c_0}$ (1 x *ns*)	vector of initial concentrations
$\mathbf{C_{in}}$ (*nf* x *ns*)	matrix of the dosing concentrations corresponding to the *nf* dosing steps
$\mathbf{C_0^{ne}}$ (*ne* x *ns*)	matrix of the varied initial concentrations corresponding to the *ne* additional experiments

Advantages of the time invariant approach:

No numerical integration required

Analytical (symbolic) relationship between the experimental conditions ($\mathbf{c_0}$, $\mathbf{C_{in}}$, $\mathbf{C_0^{ne}}$)

Mathematical description and proof: Billeter et al, Chemom. Intell. Lab. Syst., 95 (2009), 170

Strategies to treat rank deficiency in concentrations

Model reduction

Strategy 1
define *nu* uncoloured species

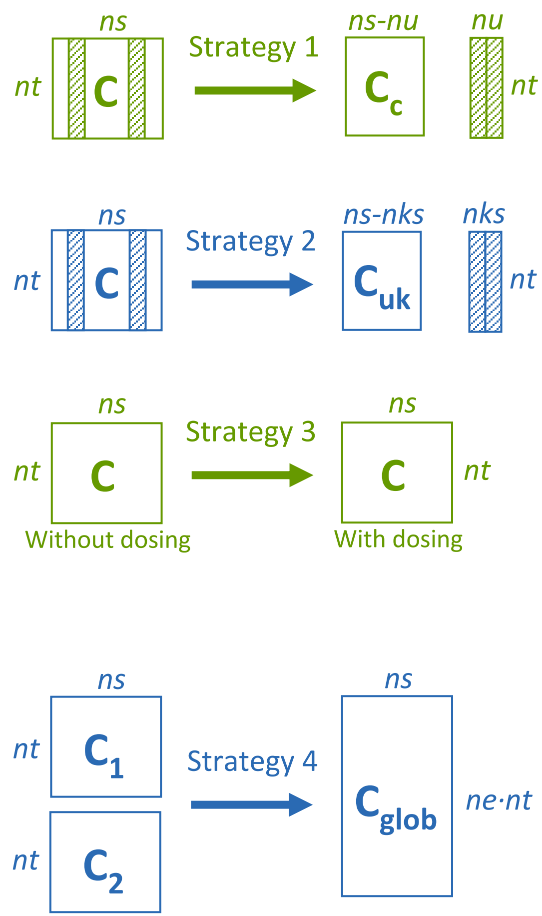
Strategy 2
include *nks* known spectra in the analysis (explicit calibration)

Rank augmentation

Strategy 3
dose one or more species in *nf* dosing steps

Strategy 4:
perform *ne* additional experiments by varying the initial concentrations (second order global analysis)

Time variant



Strategies to treat rank deficiency in concentrations

Model reduction

Strategy 1
define *nu* uncoloured species

Strategy 2
include *nks* known spectra in the analysis (explicit calibration)

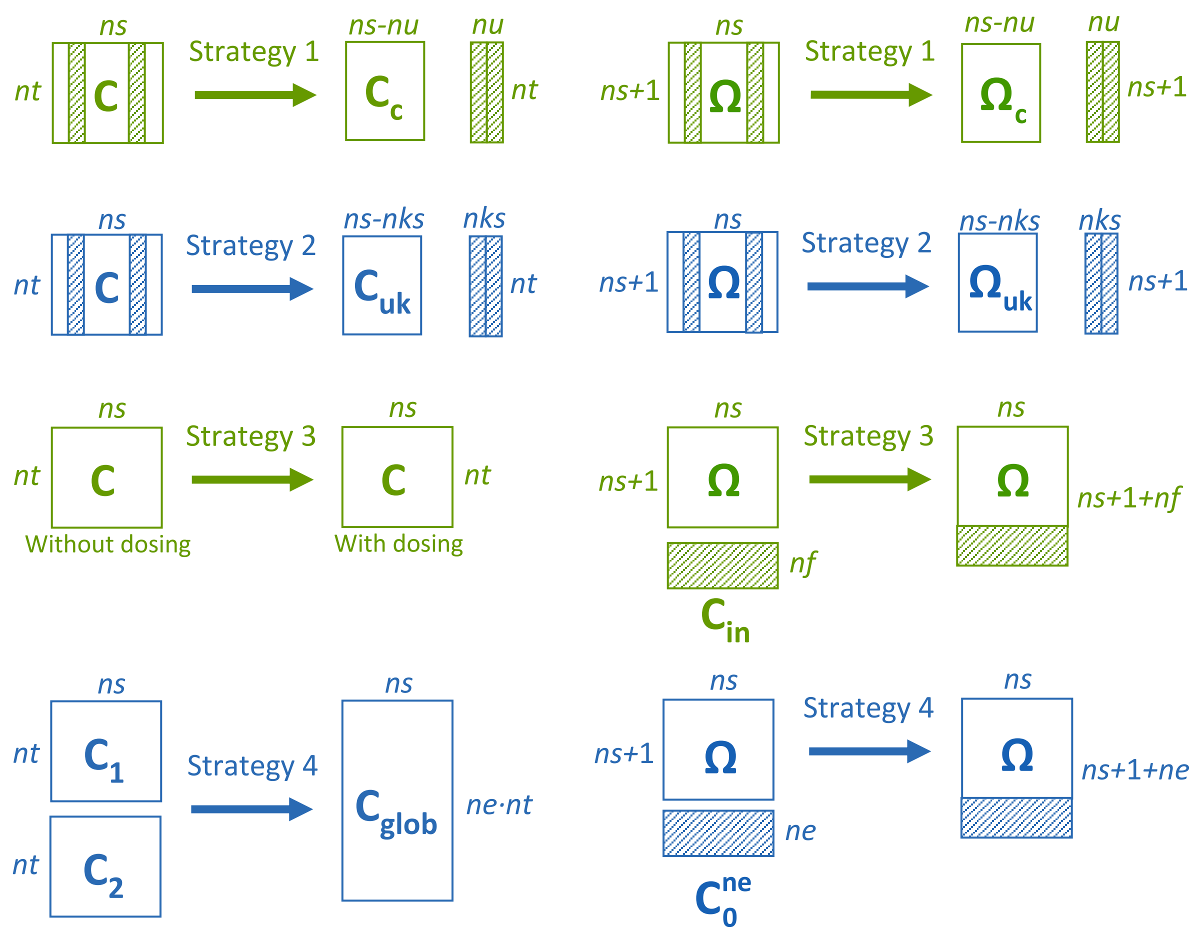
Rank augmentation

Strategy 3
dose one or more species in *nf* dosing steps

Strategy 4:
perform *ne* additional experiments by varying the initial concentrations (second order global analysis)

Time variant

Time invariant



Spectral consequence of Strategy 1

Model reduction

Strategy 1

define *nu* uncoloured species

Time variant

Time invariant

Spectral consequence of Strategy 1

Model reduction

Strategy 1
define *nu* uncoloured species

Spectral contributions of the *nu* uncoloured species are linearly transferred into the fitted pure spectra of the coloured species

The fitted component spectra $\mathbf{A}_c$ of the (*ns*–*nu*) coloured species are comprised of linear combinations of the *ns* true pure component spectra **A**

Time variant

$\mathbf{A}\ (ns \times nw)$ $\quad$ $\mathbf{A}_c\ (ns - nu \times nw)$

Time invariant

$\mathbf{A}\ (ns \times nw)$ $\quad$ $\mathbf{A}_c\ (ns - nu \times nw)$

Spectral consequence of Strategy 1

Model reduction

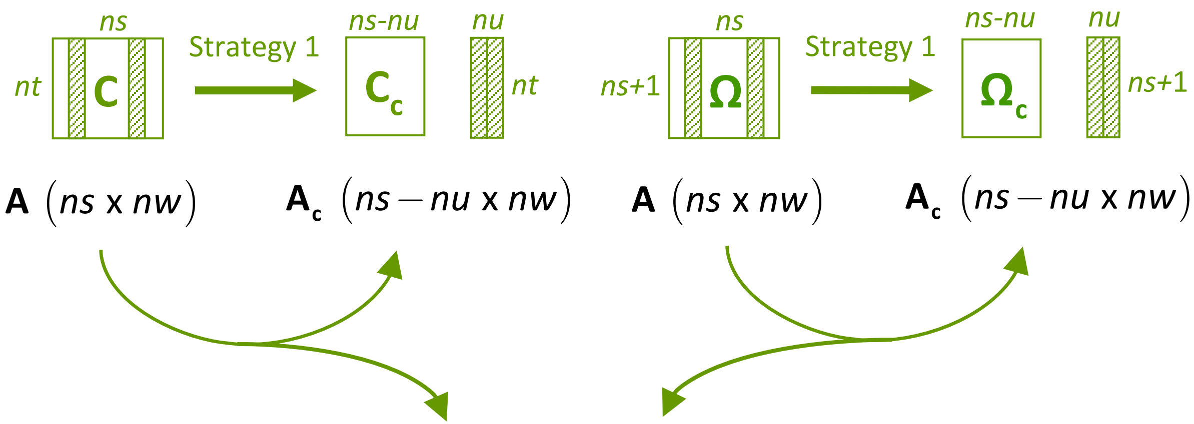
Strategy 1
define *nu* uncoloured species

Spectral contributions of the *nu* uncoloured species are linearly transferred into the fitted pure spectra of the coloured species

The fitted component spectra $\mathbf{A_c}$ of the (*ns*–*nu*) coloured species are comprised of linear combinations of the *ns* true pure component spectra $\mathbf{A}$

Time variant

Time invariant



$$\mathbf{A_c} = \boldsymbol{\Delta}\,\mathbf{A}$$

(Spectral balance)

Spectral consequence of Strategy 1

Model reduction

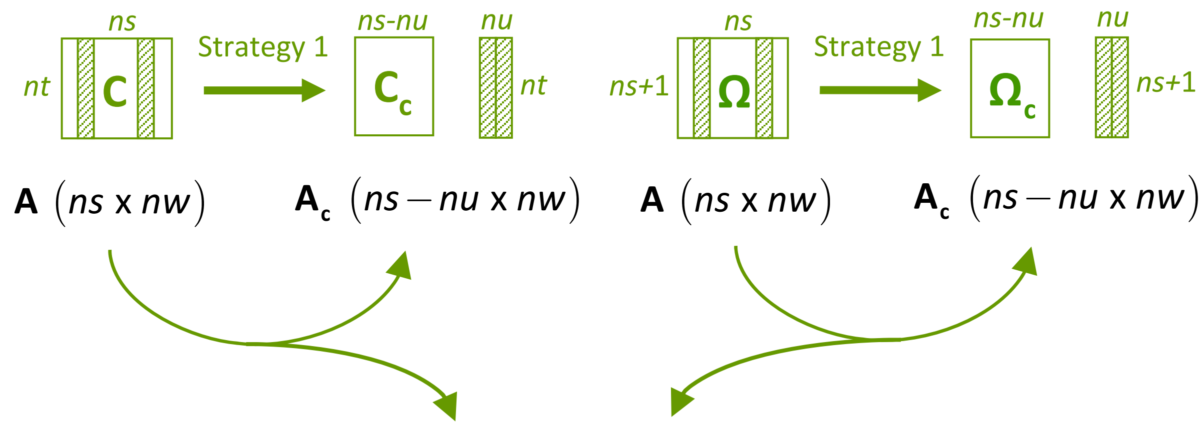
Strategy 1
define *nu* uncoloured species

Spectral contributions of the *nu* uncoloured species are linearly transferred into the fitted pure spectra of the coloured species

The fitted component spectra $\mathbf{A_c}$ of the (*ns*–*nu*) coloured species are comprised of linear combinations of the *ns* true pure component spectra **A**

Time variant

Time invariant



$$\boxed{\mathbf{A_c} = \mathbf{\Delta}\,\mathbf{A}}$$

(Spectral balance)

$$\boxed{\mathbf{\Delta}\ (ns - nu \text{ x } ns) = \mathbf{C_c^+ C} = \left(\mathbf{\Omega}\Big|_{\substack{\text{comprised of} \\ \text{coloured species}}}\right)^+ \mathbf{\Omega}\Big|_{\substack{\text{comprised of} \\ \text{all species}}}}$$

Example: Calculation of the kernel

$$A + B \xrightarrow{k_1} C$$
$$A + C \xrightarrow{k_2} D$$

Example: Calculation of the kernel

$$A + B \xrightarrow{k_1} C$$
$$A + C \xrightarrow{k_2} D$$

$$\boldsymbol{\Omega} = \left[\frac{(\mu\mathbf{1})^{\cdot\mathbf{E}^T}\ \mathrm{DIAG}(\mathbf{k})\mathbf{N}}{\mathbf{c}_0}\right] = \begin{array}{cccc} A & B & C & D \end{array}$$
$$\left[\begin{array}{cccc} -\mu k_1 - \mu k_2 & -\mu k_1 & \mu k_1 - \mu k_2 & \mu k_2 \\ -\mu k_1 - k_2 & -\mu k_1 & \mu k_1 - k_2 & k_2 \\ -k_1 - \mu k_2 & -k_1 & k_1 - \mu k_2 & \mu k_2 \\ -k_1 - k_2 & -k_1 & k_1 - k_2 & k_2 \\ \hline c_{0,A} & \alpha c_{0,A} & 0 & 0 \end{array}\right]$$

<u>Matlab code</u> (8 lines)

```
>> syms c0A alpha mu k1 k2
>> N          = [-1 -1 1 0; -1 0 -1 1];
>> E          = [1 1 0 0; 1 0 1 0];
>> k          = [k1, k2];
>> c0         = [c0A, alpha*c0A, 0, 0];
>> one        = ones(size(E'));
>> omega      = [(mu*one).^(E')*diag(k)*N; c0];
>> null(omega)
ans =
        -alpha
             1
       1-alpha
     1-2*alpha
```

$$\boxed{\alpha = \frac{c_{0,B}}{c_{0,A}}}$$

Example: Calculation of the kernel

$$A + B \xrightarrow{k_1} C$$
$$A + C \xrightarrow{k_2} D$$

$$\boldsymbol{\Omega} = \left[\frac{(\mu \mathbf{1})^{\bullet \mathbf{E}^T} \, \mathrm{DIAG}(\mathbf{k}) \mathbf{N}}{\mathbf{c}_0} \right] = \begin{array}{c} \begin{array}{cccc} A & B & C & D \end{array} \\ \left[\begin{array}{cccc} -\mu k_1 - \mu k_2 & -\mu k_1 & \mu k_1 - \mu k_2 & \mu k_2 \\ -\mu k_1 - k_2 & -\mu k_1 & \mu k_1 - k_2 & k_2 \\ -k_1 - \mu k_2 & -k_1 & k_1 - \mu k_2 & \mu k_2 \\ -k_1 - k_2 & -k_1 & k_1 - k_2 & k_2 \\ \hline c_{0,A} & \alpha c_{0,A} & 0 & 0 \end{array} \right] \end{array}$$

$$\ker \boldsymbol{\Omega} = \begin{array}{cc} \left[\begin{array}{c} -\alpha \\ 1 \\ 1-\alpha \\ 1-2\alpha \end{array} \right] & \begin{array}{c} A \\ B \\ C \\ D \end{array} \end{array}$$

Matlab code (8 lines)

```
>> syms c0A alpha mu k1 k2
>> N            = [-1 -1 1 0; -1 0 -1 1];
>> E            = [1 1 0 0; 1 0 1 0];
>> k            = [k1, k2];
>> c0           = [c0A, alpha*c0A, 0, 0];
>> one          = ones(size(E'));
>> omega        = [(mu*one).^(E')*diag(k)*N; c0];
>> null(omega)
ans =
        -alpha
             1
       1-alpha
     1-2*alpha
```

$$\boxed{\alpha = \frac{c_{0,B}}{c_{0,A}}}$$

Example: Calculation of the kernel

$$A + B \xrightarrow{k_1} C$$
$$A + C \xrightarrow{k_2} D$$

$$\mathbf{\Omega} = \left[\frac{(\mu\mathbf{1})^{\cdot\mathbf{E}^T}\,\mathrm{DIAG}(\mathbf{k})\mathbf{N}}{\mathbf{c}_0}\right] = \begin{array}{c} \begin{array}{cccc} A & B & C & D \end{array} \\ \left[\begin{array}{cccc} -\mu k_1 - \mu k_2 & -\mu k_1 & \mu k_1 - \mu k_2 & \mu k_2 \\ -\mu k_1 - k_2 & -\mu k_1 & \mu k_1 - k_2 & k_2 \\ -k_1 - \mu k_2 & -k_1 & k_1 - \mu k_2 & \mu k_2 \\ -k_1 - k_2 & -k_1 & k_1 - k_2 & k_2 \\ \hline c_{0,A} & \alpha c_{0,A} & 0 & 0 \end{array}\right] \end{array}$$

$$\ker\,\mathbf{\Omega} = \begin{array}{cc} \left[\begin{array}{c} -\alpha \\ 1 \\ 1-\alpha \\ 1-2\alpha \end{array}\right] & \begin{array}{c} A \\ B \\ C \\ D \end{array} \end{array}$$

Matlab code (8 lines)

```
>> syms c0A alpha mu k1 k2
>> N            = [-1 -1 1 0; -1 0 -1 1];
>> E            = [1 1 0 0; 1 0 1 0];
>> k            = [k1, k2];
>> c0           = [c0A, alpha*c0A, 0, 0];
>> one          = ones(size(E'));
>> omega        = [(mu*one).^(E‘)*diag(k)*N; c0];
>> null(omega)
ans =
        -alpha
             1
       1-alpha
     1-2*alpha
```

$$\boxed{\alpha = \frac{c_{0,B}}{c_{0,A}}}$$

Determination of the roots of the kernel

If $\alpha = 1$, species *C* is linearly independent from the others

If $\alpha = 0.5$, species *D* is linearly independent from the others

If $\alpha \neq 1$ or 0.5, all species are linearly dependent from the others

Example: Design of experiments

$$A + B \xrightarrow{k_1} C$$
$$A + C \xrightarrow{k_2} D$$

$$\text{ker } \mathbf{\Omega}\ (4 \times 1) = \begin{array}{cccc} A & B & C & D \\ -\alpha & 1 & 1-\alpha & 1-2\alpha \end{array}^{\mathrm{T}}$$

$\dim(\text{ker } \mathbf{\Omega}) = 1$, i.e. only one species has to be considered in Strategies (1) – (4)

Example: Design of experiments

$$A + B \xrightarrow{k_1} C$$
$$A + C \xrightarrow{k_2} D$$

$$\ker \mathbf{\Omega}\ (4 \times 1) = \begin{array}{cccc} A & B & C & D \\ \multicolumn{4}{c}{\begin{bmatrix} -\alpha & 1 & 1-\alpha & 1-2\alpha \end{bmatrix}^{\mathrm{T}}} \end{array}$$

$\dim(\ker \mathbf{\Omega}) = 1$, i.e. only one species has to be considered in Strategies (1) – (4)

Initial conditions → **Strategy** → **Species treated with current Strategy** → **Species not treated with current Strategy**

α → $\alpha = 1$ → Strategy (1) – (4) →

- *A* → *B, D,* $\boxed{C}$
- *B* → *A, D,* $\boxed{C}$
- *C* → *Not possible*
- *D* → *A, B,* $\boxed{C}$

Example: Design of experiments

$$A + B \xrightarrow{k_1} C$$
$$A + C \xrightarrow{k_2} D$$

$$\ker \boldsymbol{\Omega}\ (4 \times 1) = \begin{array}{c} \begin{array}{cccc} A & B & C & D \end{array} \\ \begin{bmatrix} -\alpha & 1 & 1-\alpha & 1-2\alpha \end{bmatrix}^{\mathrm{T}} \end{array}$$

$\dim(\ker \boldsymbol{\Omega}) = 1$, i.e. only one species has to be considered in Strategies (1) – (4)

Initial conditions	Strategy	Species treated with current Strategy	Species not treated with current Strategy
$\alpha = 1$	Strategy (1) – (4)	*A*	*B, D,* $\boxed{C}$
		B	*A, D,* $\boxed{C}$
		C	*Not possible*
		D	*A, B,* $\boxed{C}$
$\alpha = 0.5$	Strategy (1) – (4)	*A*	*B, C,* $\boxed{D}$
		B	*A, C,* $\boxed{D}$
		C	*A, B,* $\boxed{D}$
		D	*Not possible*

α

Example: Design of experiments

$$A + B \xrightarrow{k_1} C$$
$$A + C \xrightarrow{k_2} D$$

$$\ker \mathbf{\Omega}\ (4 \times 1) = \begin{array}{cccc} A & B & C & D \\ \left[-\alpha\right. & 1 & 1-\alpha & \left.1-2\alpha\right]^{\mathrm{T}} \end{array}$$

$\dim(\ker \mathbf{\Omega}) = 1$, i.e. only one species has to be considered in Strategies (1) – (4)

α — Initial conditions	Strategy	Species treated with current Strategy	Species not treated with current Strategy
$\alpha = 1$	Strategy (1) – (4)	*A*	*B, D,* $\boxed{\boldsymbol{C}}$
		B	*A, D,* $\boxed{\boldsymbol{C}}$
		C	*Not possible*
		D	*A, B,* $\boxed{\boldsymbol{C}}$
$\alpha = 0.5$	Strategy (1) – (4)	*A*	*B, C,* $\boxed{\boldsymbol{D}}$
		B	*A, C,* $\boxed{\boldsymbol{D}}$
		C	*A, B,* $\boxed{\boldsymbol{D}}$
		D	*Not possible*
$\alpha \neq 1$ or 0.5	Strategy (1) – (4)	*A*	*B,* $\boxed{\boldsymbol{C, D}}$
		B	*A,* $\boxed{\boldsymbol{C, D}}$
		C	*A, B,* $\boxed{\boldsymbol{D}}$
		D	*A, B,* $\boxed{\boldsymbol{C}}$

Example: Spectral consequence of Strategy 1

$$A + B \xrightarrow{k_1} C$$
$$A + C \xrightarrow{k_2} D$$

Model reduction

Strategy 1
define *nu* uncoloured species

Time variant

Time invariant

Example: Spectral consequence of Strategy 1

$$A + B \xrightarrow{k_1} C$$
$$A + C \xrightarrow{k_2} D$$

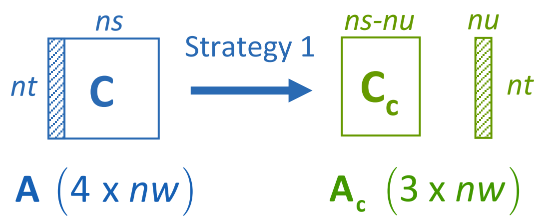
Model reduction

Strategy 1

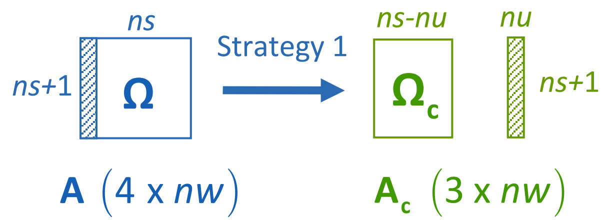
define *nu* uncoloured species

Species *A* uncoloured

Time variant



Time invariant



Example: Spectral consequence of Strategy 1

$$A + B \xrightarrow{k_1} C$$
$$A + C \xrightarrow{k_2} D$$

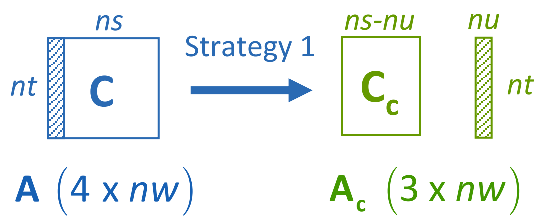
Model reduction

Strategy 1

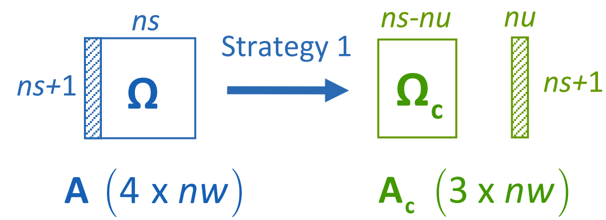
define *nu* uncoloured species

Species *A* uncoloured

Time variant



Time invariant



All species

$$\left.\boldsymbol{\Omega}\right|_{\substack{\text{comprised of}\\ \text{all species}}} = \begin{array}{cccc} A & B & C & D \\ \end{array}$$
$$\left.\boldsymbol{\Omega}\right|_{\substack{\text{comprised of}\\ \text{all species}}} = \begin{bmatrix} -\mu k_1 - \mu k_2 & -\mu k_1 & \mu k_1 - \mu k_2 & \mu k_2 \\ -\mu k_1 - k_2 & -\mu k_1 & \mu k_1 - k_2 & k_2 \\ -k_1 - \mu k_2 & -k_1 & k_1 - \mu k_2 & \mu k_2 \\ -k_1 - k_2 & -k_1 & k_1 - k_2 & k_2 \\ \hline c_{0,A} & \alpha c_{0,A} & 0 & 0 \end{bmatrix}$$

Coloured species

$$\begin{array}{ccc} 'B' & 'C' & 'D' \end{array}$$
$$\left.\boldsymbol{\Omega}\right|_{\substack{\text{comprised of}\\ \text{coloured species}}} = \begin{bmatrix} -\mu k_1 & \mu k_1 - \mu k_2 & \mu k_2 \\ -\mu k_1 & \mu k_1 - k_2 & k_2 \\ -k_1 & k_1 - \mu k_2 & \mu k_2 \\ -k_1 & k_1 - k_2 & k_2 \\ \hline \alpha c_{0,A} & 0 & 0 \end{bmatrix}$$

$$\begin{bmatrix} \mathbf{a}_{c\,'B',:} \\ \mathbf{a}_{c\,'C',:} \\ \mathbf{a}_{c\,'D',:} \end{bmatrix} = \boldsymbol{\Delta}\,\mathbf{A} = \left(\mathbf{C}\mathbf{C}^{+}\right)\mathbf{A} = \left(\left(\left.\boldsymbol{\Omega}\right|_{\substack{\text{comprised of}\\ \text{coloured species}}}\right)^{+} \left.\boldsymbol{\Omega}\right|_{\substack{\text{comprised of}\\ \text{all species}}}\right)\mathbf{A} = \begin{bmatrix} \alpha^{-1} & 1 & 0 & 0 \\ \alpha^{-1} - 1 & 0 & 1 & 0 \\ \alpha^{-1} - 2 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \mathbf{a}_{A,:} \\ \mathbf{a}_{B,:} \\ \mathbf{a}_{C,:} \\ \mathbf{a}_{D,:} \end{bmatrix}$$

Example: Spectral consequence of Strategy 1

$$A + B \xrightarrow{k_1} C$$
$$A + C \xrightarrow{k_2} D$$

Model reduction

Strategy 1
define *nu* uncoloured species

Species *A* uncoloured

Time variant

$\mathbf{A}\ (4 \text{ x } nw)$ $\quad \mathbf{A_c}\ (3 \text{ x } nw)$

Time invariant

$\mathbf{A}\ (4 \text{ x } nw)$ $\quad \mathbf{A_c}\ (3 \text{ x } nw)$

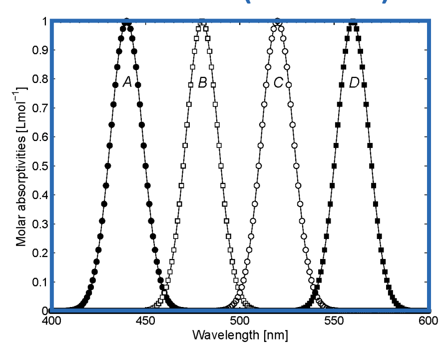
All species

$$\mathbf{\Omega}\Big|_{\substack{\text{comprised of}\\ \text{all species}}} = \begin{array}{c} \begin{array}{cccc} A & B & C & D \end{array} \\ \left[\begin{array}{cccc} -\mu k_1 - \mu k_2 & -\mu k_1 & \mu k_1 - \mu k_2 & \mu k_2 \\ -\mu k_1 - k_2 & -\mu k_1 & \mu k_1 - k_2 & k_2 \\ -k_1 - \mu k_2 & -k_1 & k_1 - \mu k_2 & \mu k_2 \\ -k_1 - k_2 & -k_1 & k_1 - k_2 & k_2 \\ \hline c_{0,A} & \alpha c_{0,A} & 0 & 0 \end{array}\right] \end{array}$$

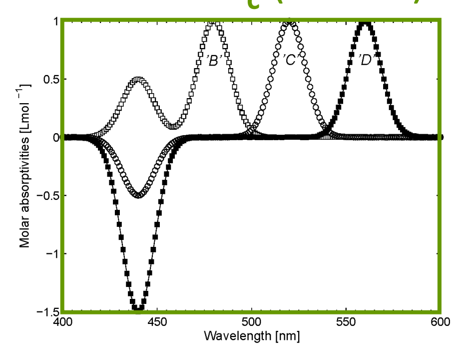
Coloured species

$$\mathbf{\Omega}\Big|_{\substack{\text{comprised of}\\ \text{coloured species}}} = \begin{array}{c} \begin{array}{ccc} 'B' & 'C' & 'D' \end{array} \\ \left[\begin{array}{ccc} -\mu k_1 & \mu k_1 - \mu k_2 & \mu k_2 \\ -\mu k_1 & \mu k_1 - k_2 & k_2 \\ -k_1 & k_1 - \mu k_2 & \mu k_2 \\ -k_1 & k_1 - k_2 & k_2 \\ \hline \alpha c_{0,A} & 0 & 0 \end{array}\right] \end{array}$$

True: **A** (4 x *nw*)



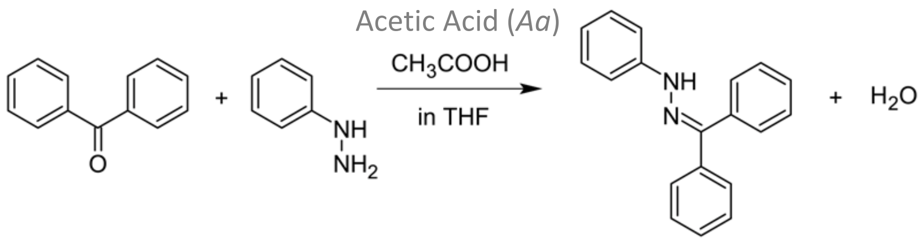
Fitted: $\mathbf{A_c}$ (3 x *nw*)



$$\begin{bmatrix} \mathbf{a}_{c\,'B',:} \\ \mathbf{a}_{c\,'C',:} \\ \mathbf{a}_{c\,'D',:} \end{bmatrix} = \mathbf{\Delta}\,\mathbf{A} = \left(\mathbf{C}\mathbf{C}^{+}\right)\mathbf{A} = \left(\left(\mathbf{\Omega}\Big|_{\substack{\text{comprised of}\\ \text{coloured species}}}\right)^{+} \mathbf{\Omega}\Big|_{\substack{\text{comprised of}\\ \text{all species}}}\right)\mathbf{A} = \begin{bmatrix} \alpha^{-1} & 1 & 0 & 0 \\ \alpha^{-1}-1 & 0 & 1 & 0 \\ \alpha^{-1}-2 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \mathbf{a}_{A,:} \\ \mathbf{a}_{B,:} \\ \mathbf{a}_{C,:} \\ \mathbf{a}_{D,:} \end{bmatrix} = \begin{bmatrix} 0.5 \\ -0.5 \\ -1.5 \end{bmatrix} \mathbf{a}_{A,:} + \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \mathbf{a}_{B,:} + \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \mathbf{a}_{C,:} + \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \mathbf{a}_{D,:} \quad \text{when } \alpha = 2$$

An experimental case study

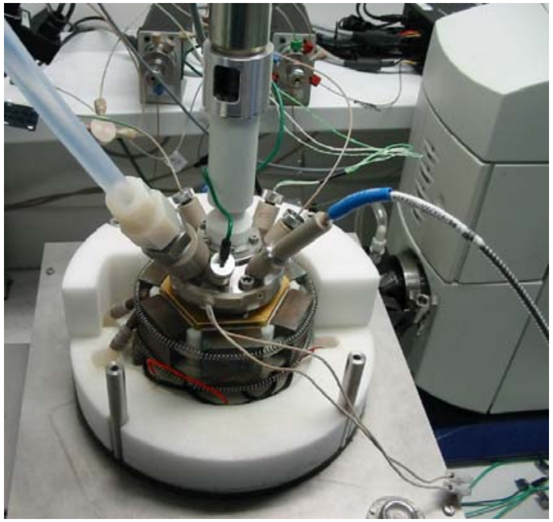
Overall reaction



Benzophenone (*B*)

Phenylhydrazine (*P*)

Benzophenone-Phenylhydrazone (*BP*)



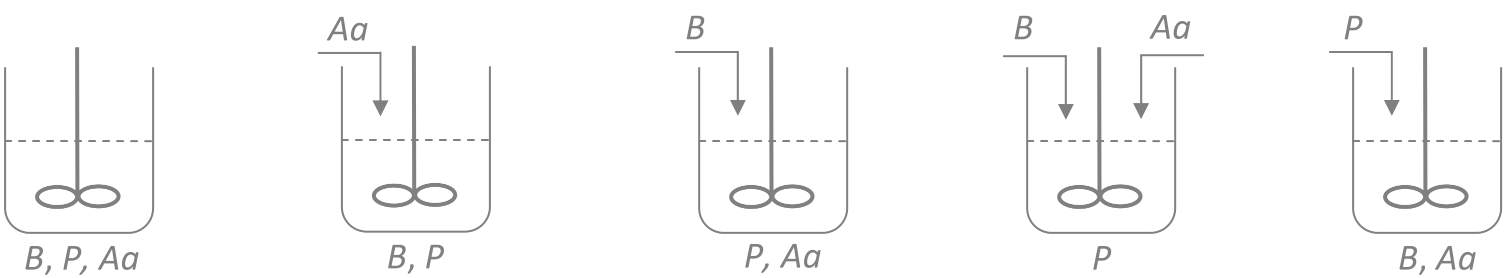
Kinetic mechanism

$$B + P + Aa \xrightarrow{k} BP + Aa$$

Reactor: CRC.v4 with FT-IR and UV-vis

Experimental conditions

25°C, mid-IR (1200–1650 cm^{-1}), UV-vis (240–400 nm)

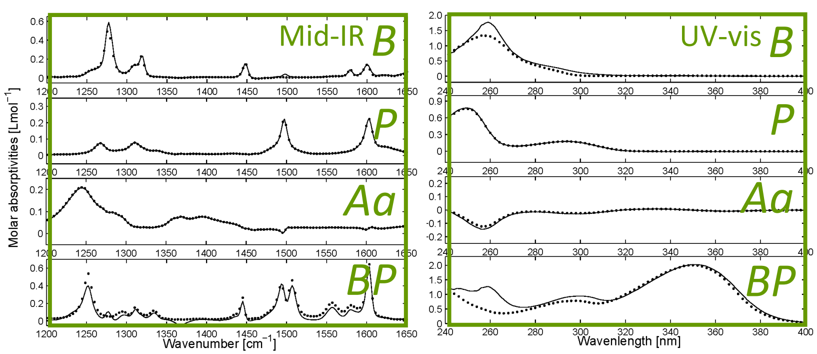


Batch conditions | **Dosing *Aa*** | **Dosing *B*** | **Dosing *B* + *Aa*** | **Dosing *P***

Spectral validation

$$B + P + Aa \xrightarrow{k_1} BP + Aa$$

Strategy (3): dosing



Species *B* and *Aa* dosed

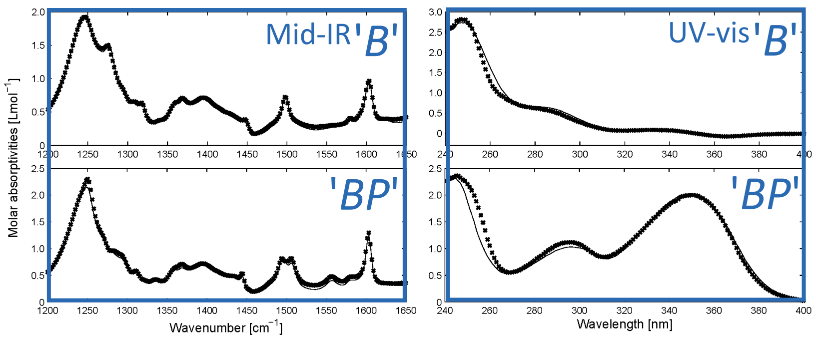
Full spectral resolution

—— Fitted spectra ••• Measured spectra **XXX** Predicted spectra

Spectral validation

$$B + P + Aa \xrightarrow{k_1} BP + Aa$$

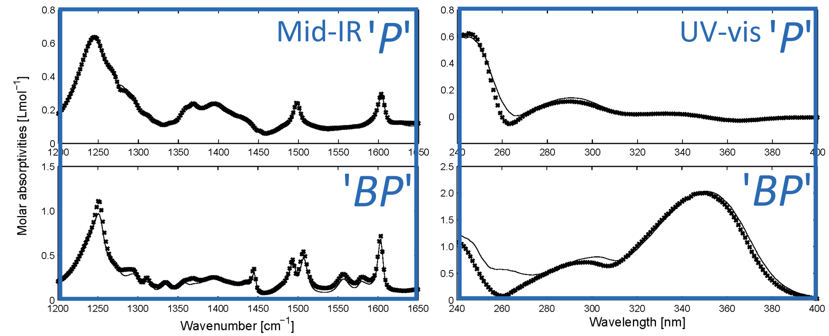
Strategy (1): uncoloured species



Species *P* and *Aa* set uncoloured

Partial spectral resolution

Strategy (1)+(2): provided known spectrum



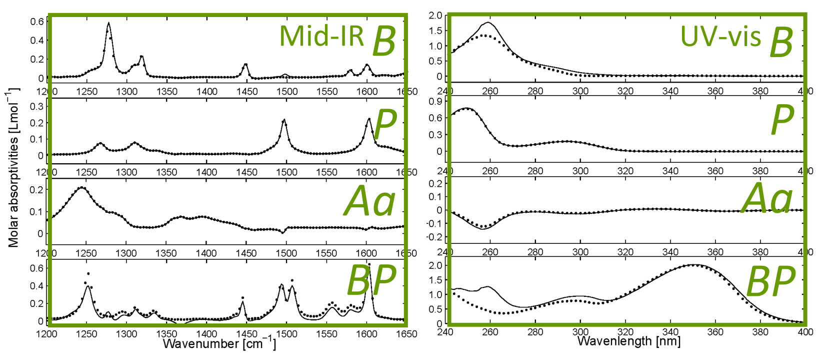
Pure spectrum of *B* provided

Aa set uncoloured

Partial spectral resolution

The model can now be validated **with** and **without** rank deficiency!

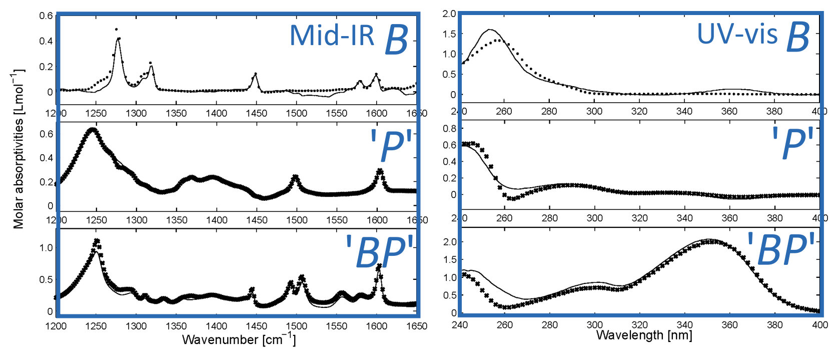
Strategy (3): dosing



Species *B* and *Aa* dosed

Full spectral resolution

Strategy (1)+(4): second order global analysis

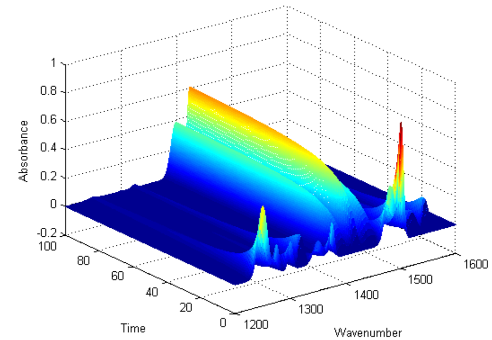


Initial concentration of *B* varied, *Aa* set uncoloured

Partial spectral resolution

—— Fitted spectra ●●● Measured spectra **XXX** Predicted spectra

Conclusions



$$\begin{cases} A + B \xrightarrow{k_1} C \\ A + C \xrightarrow{k_2} D \end{cases}$$

Ker **C**

Linear dependencies, leading to rank deficiency in concentration matrix, can be elucidated using the **concept of kernel**

C(ker **C**) = **0**

This equation defines a **mass balance equation** with time-invariant coefficients

Model reduction / rank augmentation Strategies

Species to be included in Strategies 1 – 4 can be identified by selecting the species with **non zero rows in the kernel**

Matrix **Ω**

The kernel of the time-variant concentrations can be calculated without numerical integration using a simple **time-invariant approach**

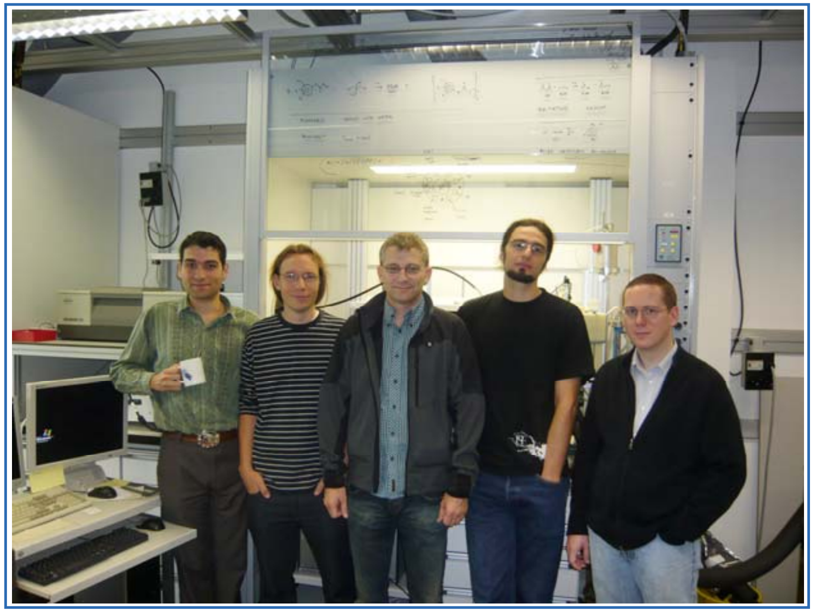
Matrix **Δ**

Linear combinations, observed in the fitted pure spectra when Strategy 1 (defining uncoloured species) is used, can be explained by a **spectral balance equation**

Spectral validation of kinetic models

Spectral validation of kinetic models is now possible even in case of **rank deficient spectroscopic data**

Reaction Analysis at ETH Zürich



Reaction Analysis Group
Leader: Dr. Bobby Neuhold

Dr. Bobby Neuhold, Chemist
Chemometrics, Kinetic hard-modelling and Spectroscopy

Dr. Gilles Richner, Chemical Engineer
Calorimetry and Hydrogenation

Julien Billeter, Chemical Engineer
Chemometrics, Kinetic hard-modelling and Spectroscopy

Tamás Godány, Chemical Engineer
Low-temperature alkylations

Sébastien Cap, Chemist
Particulate systems and Dissolution

Safety and Environmental Technology Group

under supervision of **Prof. K. Hungerbühler**

ETH
Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

Thank you for your attention

Publications

Systematic prediction of linear dependencies in the concentration profiles and implications on the kinetic hard-modelling of spectroscopic data
Billeter et al, Chemom. Intell. Lab. Syst., 95 (2009), 170 – 187

Kinetic hard-modelling and spectral validation of rank-deficient spectroscopic data: a case study
Billeter et al, Chemom. Intell. Lab. Syst., submitted
